Available online at www.sciencedirect.com

Journal of Applied Research and Technology

Journal of Applied Research and Technology 14 (2016) 338–344

www.jart.ccadet.unam.mx

Original

# Machine Learning for benthic sand and maerl classification and coverage estimation in coastal areas around the Maltese Islands

Adam Gauci [a,b,f,*], Alan Deidun [a], John Abela [c,f], Kristian Zarb Adami [d,e,f]

[a] *Department of Geosciences, Faculty of Science, University of Malta, Malta*
[b] *Department of Intelligent Computer System, Faculty of ICT, University of Malta, Malta*
[c] *Department of Computer Information Systems, Faculty of ICT, University of Malta, Malta*
[d] *Department of Physics, Faculty of Science, University of Malta, Malta*
[e] *Department of Physics (Astrophysics), University of Oxford, UK*
[f] *Institute of Space Sciences and Astronomy, University of Malta, Malta*

Received 14 March 2016; accepted 15 August 2016
Available online 27 October 2016

**Abstract**

Analysis of the seabed composition over a large spatial scale is an interesting yet very challenging task. Apart from the field work involved, hours of video footage captured by cameras mounted on Remote Operated Vehicles (ROVs) have to be reviewed by an expert in order to classify the seabed topology and to identify potential anthropogenic impacts on sensitive benthic assemblages. Apart from being time consuming, such work is highly subjective and through visual inspection alone, a quantitative analysis is highly unlikely to be made. This study investigates the applicability of various Machine Learning techniques for the automatic classification of the seabed into maerl and sand regions from recorded ROV footage. ROV data collected from depths ranging between 50 m and 140 m and at 9.5 km from the northeast coastline of the Maltese Islands, is processed. Through the application of the presented technique, 5.23 GB of data corresponding to 2 h and 24 min of footage which was collected during June 2013, was initially cleaned and classified. An estimate for the percentage cover of the two benthic habitats (sandy seabed and maerl) was also computed by using artifacts encountered during the ROV survey and of known dimensions as a reference. Unlike other automatic seabed mapping techniques, the presented prototype processes video footage captured by a down-facing camera and not through acoustic backscatter. Image data is easier and much cheaper to capture. Promising results that indicate a very good degree of agreement between the true and predicted habitat type distribution values, were obtained.


*Keywords:* Machine Learning; Image processing; Seabed classification; Decision trees; Maerl detection; Sand detection

## 1. Introduction

In the last few decades there has been a dramatic upsurge in the laying of submarine cables and pipelines that are deployed mainly for communications or energy-transfer purposes. Such laying inevitably implies a number of environmental changes which warrant the conduct of environment impact assessment (EIA) studies. Moreover, pipelines are normally deployed at great depths, beyond safe SCUBA diving limits. Remotely operated vehicles (ROVs) are used to collect baseline data as well as to monitor the environmental changes. Whilst unaided human analysis of such video footage allows one to infer qualitative conclusions about the seabed type being studied, the conduction of quantitative assessments through manual means is virtually impossible and highly subjective.

In order to ensure a higher degree of security in energy supply, in 2013 the Maltese government reached an agreement with the Italian government to connect the islands with the European electrical grid. This involved the laying of two 95 km-long submarine electrical cables between Qalet Marku in Malta and Marina di Ragusa in Sicily. As shown by Borg,

* Corresponding author.
*E-mail address:* adam.gauci@um.edu.mt (A. Gauci).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

http://dx.doi.org/10.1016/j.jart.2016.08.003

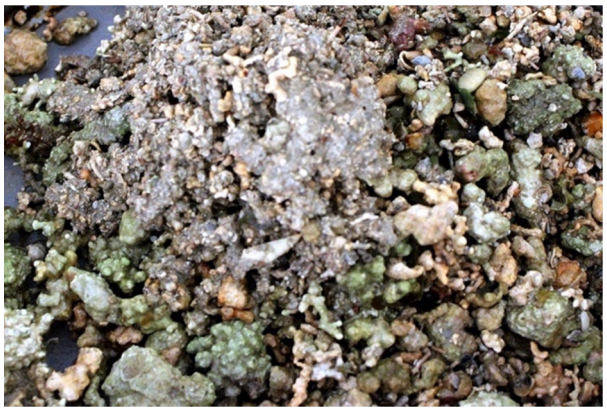

Fig. 1. Photo of a sample of maerl taken within a laboratory.

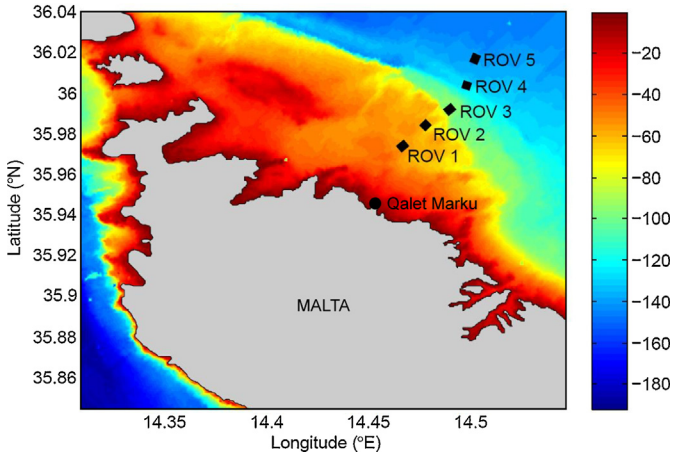


Fig. 2. Bathymetry close to the Maltese islands and the five seabed stretches of 500 m (total of 2.5 km), staggered through 1000 m-long intervals, over which ROV footage was recorded.

Lanfranco, Mifsud, Rizzo, and Schembri (1999), Dimech, Borg, and Schembri (2004), Sciberras et al. (2009), and Agnesi et al. (2009), the planned transect intersected with maerl assemblages, which consist of accumulations of calcareous rhodophyte thalli belonging mainly to Corallinaceae and marginally to Peyssonneliaceae (Fig. 1).

In this study, the applicability of various Machine Learning (ML) techniques for the automatic classification of the seabed into maerl and sand regions from recorded ROV footage, was investigated. The developed prototype contributes by providing an alternative or to supplement costly and highly labor-intensive acoustic benthic mapping techniques. The manual analyses of collected video footage is also elevated and a non-subjective approach to quantify the seabed type is made available.

Examples of artificial intelligence (AI) methods used for seabed classification can be found in literature; however, most studies consider acoustic backscatter data. Stephens and Diesing (2014) used multi-beam echo-sounder data collected from the North Sea to classify the substrate types. In this case, the ground truth was determined through the collected samples. Moškon, Žibert, and Kavšek (2015) also demonstrated how classifiers can be applied to raw multi-beam acoustic data for seabed mapping. Similar studies were carried out by Coiras and Williams (2009) and Landmark, Solberg, Austeng, and Hansen (2014). Investigations of seabed classification from image data were not found. As suggested by Stephens and Diesing (2014), reliable automated approaches that provide quantitative results and which can be included in monitoring programs are still relatively novel. While backscattered acoustic signals highlight the differences in the seabed clearly, surveys making use of such signals require a lot of planning, a lengthy permitting procedure, and are very expensive to run. In this study, camera footage captured by an ROV is used. The required data can also be recorded by a towed down-facing camera and is independent of the water transparency and of differences in the lightning conditions.

The following section provides further details on the video footage and how the frames were extracted. Details about the classification methods used and the obtained results are discussed in Sections 3 and 4, respectively. Planned future work and concluding remarks are given in Section 5.

## 2. Image data

For this study, 2 h and 24 min of video footage which were recorded by the down-facing camera of the ROV over five seabed stretches of 500 m were processed (Fig. 2). The ROV transects followed the path proposed for the submarine interconnector cable. The RAW video was captured at 29 frames per second and recorded on digital media in MPEG2 format. The processing was carried out on frames of 720 × 480 pixels. Since the velocity of the ship was kept constant during data capture, the coordinates of each frame could be computed from the corresponding timestamps. In Table 1, the starting and ending coordinates (in decimal degrees), the average depth, the average vessel speed and the number of frames extracted for every stretch are summarized.

The resolution of the raster set was estimated with frames that showed artifacts of known dimensions. In particular, reference was made to a bomb shell dating from World War II which was captured in a number of frames (Fig. 3). As documented in Boyd (2009), the body length (from the tip without the tail fins) and the diameter at the widest part of such 500-lb general purpose bombs are of 90.67 cm and 30.02 cm, respectively. The number of pixels representing these lengths were obtained from five different frames and the averages were found to be 301.84 and 92.87 pixels. By using these two measurements, an estimate for the physical dimension of individual pixels could be calculated as being 0.003004 m and 0.003233 m, with an average of 0.003118 m. Since the frames consisted of 720 × 480 pixels, each scene represented an area of 2.2451 m × 1.4968 m.

Some of the frames were found to contain periodic noise that contaminated the signal. Such frames were projected onto the Fourier domain and any high frequency components away from the central axes were masked out to remove this periodic pattern. Figure 4 depicts the noise removal process on the blue channel in both the image and Fourier spaces. Through this enhancement, no extra information or artifacts was added to the image data and the luminous values remained unchanged.

Table 1
Average depth, average vessel speed and number of frames extracted for every ROV stretch.

| ROV | Starting coordinates | Stopping coordinates | Average depth (m) | Average vessel speed (m/s) | Number of extracted frames |
|---|---|---|---|---|---|
| 1 | 35.972830° N, 14.465499° E | 35.976048° N, 14.469327° E | 53.10 | 0.2462 | 33 |
| 2 | 35.982762° N, 14.476914° E | 35.986069° N, 14.480641° E | 60.40 | 0.3910 | 99 |
| 3 | 35.990776° N, 14.489086° E | 35.994190° N, 14.492703° E | 96.84 | 0.2811 | 100 |
| 4 | 36.002065° N, 14.498090° E | 36.006408° N, 14.499593° E | 96.84 | 0.3098 | 100 |
| 5 | 36.015201° N, 14.502245° E | 36.019428° N, 14.504150° E | 138.07 | 0.3038 | 99 |

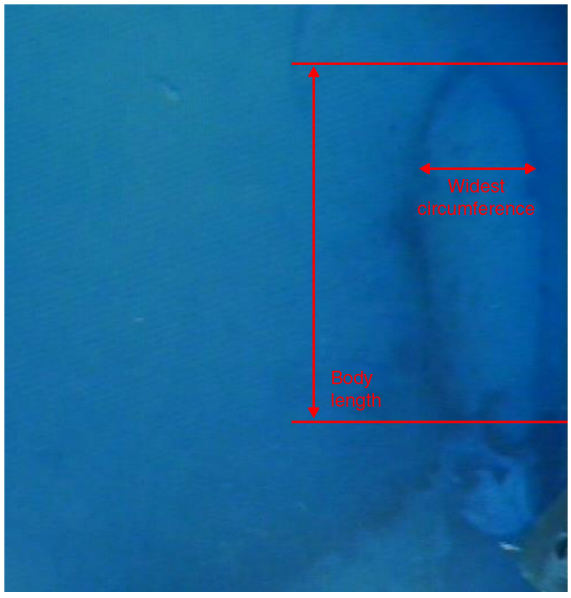


Fig. 3. WW2 bomb shell used to calibrate for size.

## 3. Feature extraction and classification methods

The task of seabed classifying is a typical classification problem in which samples or pixels are to be categorized into groups. The main objective of this study was to investigate which model most accurately predicts whether a pixel belongs to a maerl or sand patch. All classification methods were tested using tuples storing RGB or LAB information to verify whether the classification performance increases when representing the data in different color spaces. Tuples encoding RGB data stored three numbers that ranged from 0 to 255 and which represented the red, the green and the blue intensities of each pixel. When data was projected onto the LAB model, the L channel represented the luminance, the A channel showed the variation from green to red, while the B channel showed the variation from blue to yellow. Typical training examples and the corresponding values are shown in Figure 5.

Decision tree learning is based on the pioneering work done by Hunt in the late 1950s. Early in the 1960s, Quinlan developed the Iterative Dichottomizer 3 (ID3). This was followed by the improved C4.5 decision tree learners (Kohavi & Quinlan, 2002). In this work, the J48 method (an implementation of the

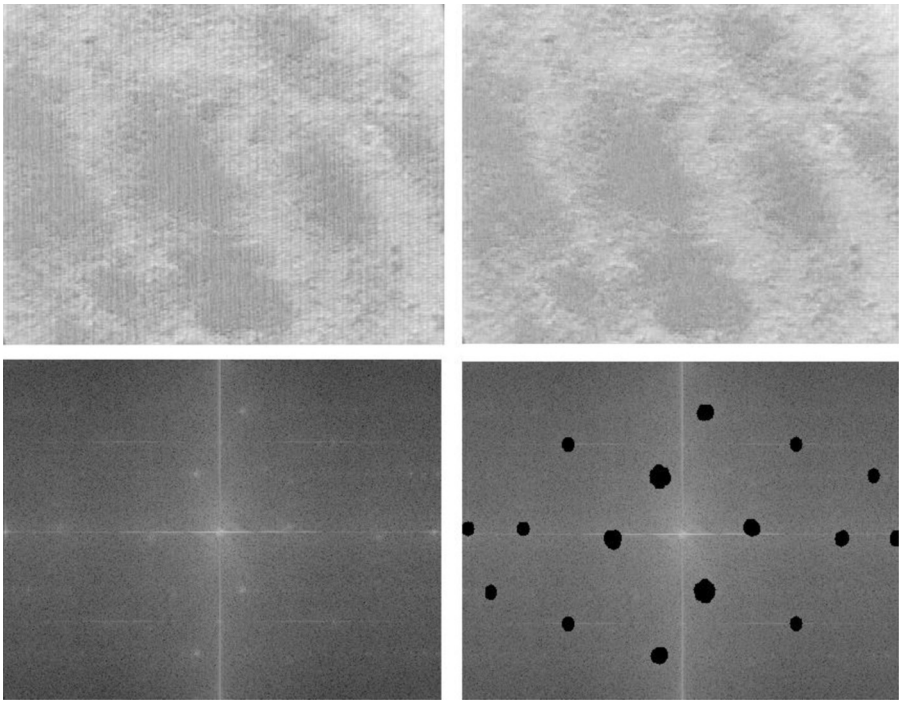

Fig. 4. Original frame (left) and the cleaned version (right) in image space (top) and Fourier space (bottom).

| Training example | R | G | | L | A | B | Class | Colour sample |
|---|---|---|---|---|---|---|---|---|
| 1 | 72 | 147 | 136 | 143 | 101 | 126 | 1 | |
| 2 | 12 | 114 | 166 | 114 | 116 | 92 | 2 | |
| 3 | 88 | 146 | 136 | 144 | 106 | 127 | 1 | |
| 4 | 32 | 131 | 152 | 128 | 105 | 108 | 2 | |
| 5 | 84 | 147 | 139 | 144 | 105 | 126 | 1 | |

Fig. 5. Training sample representations in RGB and CieLAB spaces together with the corresponding class number and the color.

C4.5 algorithm), the CART (Classification and Regression Tree) and Random Forests (with 10 as well as 50 trees) methods, were tested. Such classification schemes sort samples by determining the corresponding leaf node after traversing down the data structure from the root. Most of the available methods construct the tree by adopting a greedy search strategy and use an information gain evaluation function to determine whether or not an attribute can represent the training samples. Branches of the tree are built by recursively repeating this process for each node and the process stops when all elements in the subset at a point have the same value as the target variable, or when splitting no longer adds value to the predictions (Mitchell, 1997). The constructed tree can be then used as a rule set for predicting the category of an unknown sample from the same set of attributes.

In this study, the applicability of an artificial neural network (ANN) for the classification of data in maerl and sand pixels, was also investigated. Such a learning model connects a number of neurons that take a set of inputs and produce a single real number. The learning algorithm determines numerical weights to apply between each of these elements to obtain the desired output. An advantage of this technique is that it can produce good results even when the input is noisy or incomplete. In particular, a backward propagation network with a 3:7:1 structure was constructed. In Figure 6, combinations of RGB and LAB channels are visualized to highlight natural separation of the two classes. These graphs show that even in two dimensions, the data is separable and hence the neural network model was expected to be able to accurately learn how to differentiate between sand and maerl pixels.

## 4. Results and model performance

The performance of each model was tested on separate data sets for each individual ROV stretch, as well as on a combined and randomized global set. Classification was carried out on RGB and LAB spaces. The labeled data was manually created by importing the frames in Adobe Photoshop and painting over maerl and sand patches by solid green and solid blue colors, respectively. The created masks were then imported in Matlab to extract the class information (Fig. 7).

Testing was performed on labeled data using 10 fold cross-validation. In each experiment, the data was divided into ten

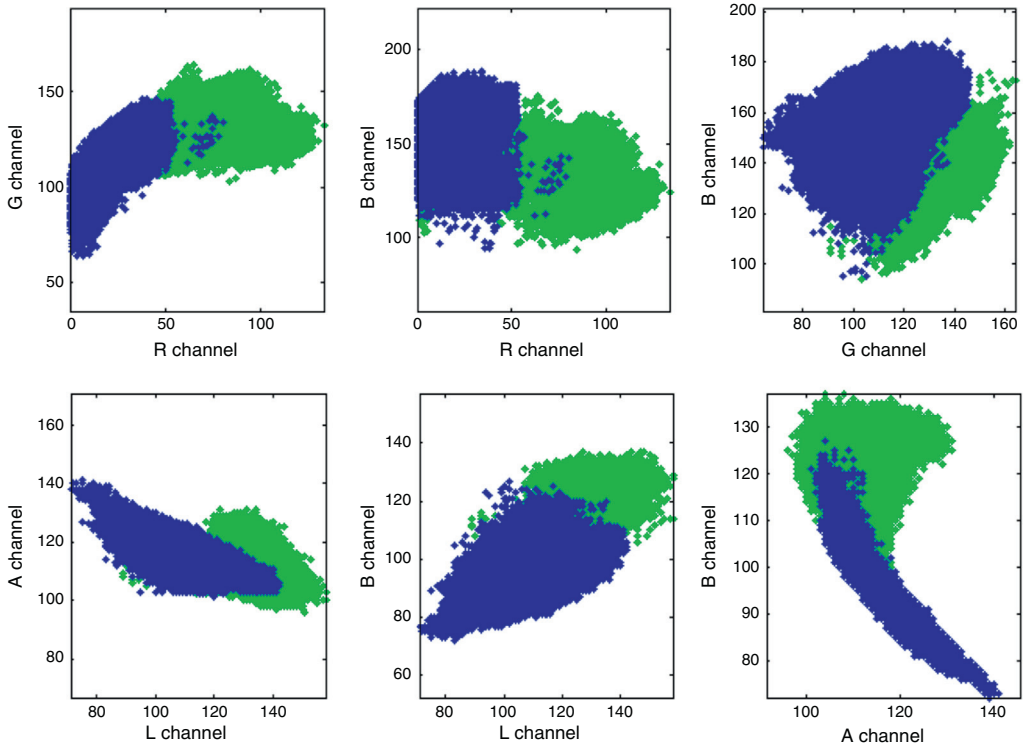


Fig. 6. Combinations of RGB color channels (top) and CieLAB color channels (bottom) to show natural separation between sand pixels (green) and maerl pixels (blue).

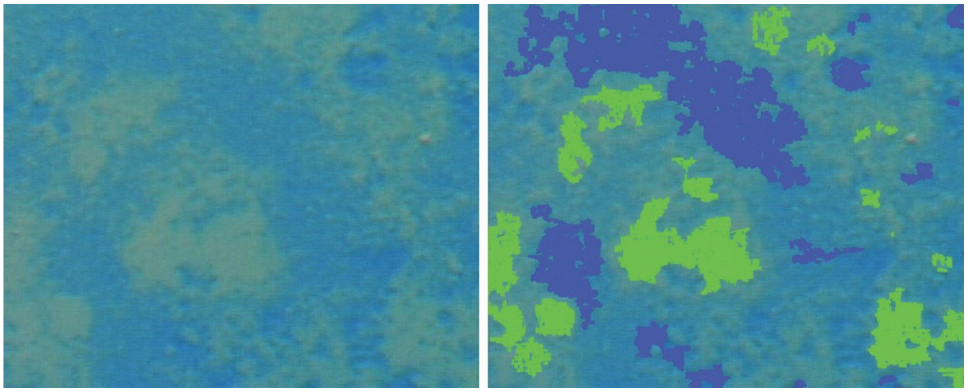

Fig. 7. Original frame from which training values are extracted (left) and the manually generated mask to highlight the pixel locations to be used for training and the corresponding taxonomy class (right).

groups and the algorithms were run for ten times. At every iteration, nine groups were used for training while the remaining group was used to test the performance of the model. Such a methodology ensured that each training sample was used for testing at least once.

The resulting confusion matrices that show the number of correctly and incorrectly classified samples for all tested methods, are presented in Figure 8. In Figure 9, the actual percentage accuracies which represent the ratio between the sum of true positives and true negatives are listed together with the total number of samples processed. Following the classification of every pixel, an estimation of the total benthic area covered was determined. The results for the first three ROV stretches are presented in Table 2.

In this study, all ML algorithms produced good results. For half of the tested cases, the J48 algorithm produced the best results. The CART method gave the best performance in the remaining cases. While good results were also achieved by Random Forest with 10 trees, significantly better accuracies were not recorded when using 50 trees. The additional time and memory

| | | ROV 1 | | ROV 2 | | ROV 3 | | ROV All | |
|---|---|---|---|---|---|---|---|---|---|
| J48 | RGB | 229 773 | 75 | 48 974 | 81 | 94 796 | 728 | 373 275 | 1152 |
| | | 170 | 68 471 | 44 | 50 559 | 75 | 123 182 | 642 | 241 859 |
| | LAB | 229 713 | 135 | 48 956 | 99 | 94 705 | 819 | 373 176 | 1251 |
| | | 151 | 68 490 | 44 | 50 559 | 186 | 123 071 | 871 | 241 630 |
| Cart | RGB | 229 774 | 74 | 48 964 | 91 | 94 819 | 705 | 373 309 | 118 |
| | | 50 | 68 491 | 37 | 50 566 | 100 | 123 157 | 676 | 241 825 |
| | LAB | 229 718 | 130 | 48 954 | 101 | 94 713 | 811 | 373 162 | 1265 |
| | | 152 | 68 489 | 58 | 50 545 | 195 | 123 062 | 857 | 241 644 |
| Random forest (10 trees) | RGB | 229 731 | 117 | 48 974 | 81 | 94 903 | 621 | 373 272 | 1155 |
| | | 151 | 68 490 | 73 | 50 530 | 333 | 122 924 | 934 | 241 567 |
| | LAB | 229 726 | 122 | 48 965 | 90 | 94 743 | 781 | 373 158 | 1269 |
| | | 177 | 68 464 | 69 | 50 534 | 235 | 123 022 | 867 | 241 634 |
| Random forest (50 trees) | RGB | 229 732 | 116 | 48 967 | 88 | 94 896 | 628 | 373 254 | 1173 |
| | | 148 | 68 493 | 66 | 50 537 | 285 | 122 972 | 896 | 241 605 |
| | LAB | 229 730 | 118 | 48 965 | 90 | 94 743 | 781 | 373 160 | 1267 |
| | | 177 | 68 464 | 65 | 50 538 | 221 | 123 036 | 850 | 241 651 |
| Neural network | RGB | 229 751 | 97 | 48 966 | 89 | 94 801 | 723 | 373 302 | 1125 |
| | | 149 | 68 492 | 38 | 50 565 | 187 | 123 070 | 769 | 241 732 |
| | LAB | 229 734 | 114 | 48 955 | 100 | 94 673 | 851 | 373 041 | 1386 |
| | | 214 | 68 427 | 34 | 50 569 | 146 | 123 111 | 743 | 241 758 |

Fig. 8. Confusion matrices for all runs performed. Each 2 × 2 grid shows the true positive, false negative, false positive and true negative in row major order.

| | | ROV 1 | ROV 2 | ROV 3 | ROV All |
|---|---|---|---|---|---|
| J48 | RGB | 0.99918 | 0.99875 | 0.99633 | 0.99709 |
| | LAB | 0.99904 | 0.99857 | 0.99541 | 0.99656 |
| Cart | RGB | 0.99958 | 0.99872 | 0.99632 | 0.99871 |
| | LAB | 0.99906 | 0.99840 | 0.99540 | 0.99656 |
| Random forest (10 trees) | RGB | 0.99910 | 0.99845 | 0.99564 | 0.99661 |
| | LAB | 0.99900 | 0.99840 | 0.99536 | 0.99654 |
| Random forest (50 trees) | RGB | 0.99912 | 0.99845 | 0.99583 | 0.99665 |
| | LAB | 0.99901 | 0.99844 | 0.99542 | 0.99657 |
| Neural network | RGB | 0.99918 | 0.99873 | 0.99584 | 0.99693 |
| | LAB | 0.99890 | 0.99866 | 0.99544 | 0.99655 |

Fig. 9. Percentage accuracies for all runs performed.

Table 2
Results for the first three ROV stretches.

| | ROV 1 | ROV 2 | ROV 3 |
|---|---|---|---|
| Number of valid frames | 32 | 99 | 6 |
| Total number of pixels processed | 9,574,400 | 29,478,339 | 1,786,566 |
| Pixels classified as sand | 5,493,067 (57.37%) | 10,314,343 (34.99%) | 876,397 (49.05%) |
| Pixels classified as maerl | 4,081,333 (42.63%) | 19,163,996 (65.01%) | 910,169 (50.95%) |
| Estimated total surface area coverage ($m^2$) | 93.0816 | 286.5862 | 17.5000 |
| Estimated sand surface area coverage ($m^2$) | 53.4032 | 100.2753 | 8.5203 |
| Estimated maerl surface area coverage ($m^2$) | 39.6784 | 186.3000 | 8.8000 |

required to carry out such runs is not worth the small ($\times 10^{-5}$) increase in percentage accuracy. The same applies for predictions made with the ANN method.

The J48 and CART methods constructed a representative decision tree in a few seconds. The generated data structure could be easily parsed to provide an accurate output in a very short time. Such techniques can only be used when the output class is discretized; nevertheless, most real value systems can be adapted to work with data bins that range over a very small interval, allowing for these classification methods to be utilized.

In this study, experts in seabed morphology classification were involved at all the stages so as to ensure that the model outputs made sense and agree with what is expected in the real world. Such techniques can easily be applied by other scientists working in other fields of research, such as terrestrial vegetation classification or in the generation of cloud cover masks.

## 5. Conclusion

This work investigated various Machine Learning techniques for automated seabed classification into maerl and sand patches. The very good and promising results that were obtained suggest that the proposed methods can be used as an alternative or to supplement more costly or more labor-intensive acoustic benthic mapping techniques which include multibeam sonar and side-scan sonar. The laborious and time-consuming manual analyses of collected video footage or still images of the seabed is also elevated.

As summarized in the study for NOAA (Finkbeiner, Stevenson, & Seaman, 2001), imagery and remote sensing methodologies are becoming frequently used for the identification of benthic assemblages. Whilst manual examination of video footage and still images of maerl assemblages is common in the literature such as in Hinz et al. (2010), very few studies, if any, have ever proposed an automated image analysis protocol for ROV footage in order to determine benthic percentage cover of maerl. Expert manual examination of ROV footage has the benefits of enhanced taxonomic resolution over automated examination of the same material and allows for taxonomic identification down to a much lower taxonomic rank. Identification of different maerl bed types is even possible; nonetheless, manual techniques are time-intensive, require taxonomic expert input and can only provide approximate quantitative measures of seabed cover of different benthic assemblages for comparative purposes. Both manual and automated techniques should be accompanied by ground truthing and calibration through sample collection from the field for ex-situ examination, and this routine procedure could not be conducted in the present study in view of the prohibitive sea depths involved.

Apart from noise reduction and the binary classification of pixels, this work also investigated the possibility of computing estimates for surface area coverage in square meters. This was

made possible by means of detected debris with known physical dimensions as a scale.

While the results prove that very good classification results can be achieved, the quoted area estimates for maerl and sand heavily rely on the validity of the assumption that the ROV was kept at a constant height above the sea bed. Calibration for pixel dimensions was only performed from a few frames in which debris was spotted and the viewing conditions might have varied even across footage of the same ROV stretch. Although footage was across stretches of 500 m and the sea floor topology in the area being studied goes down gradually, keeping the ROV at a constant altitude is very difficult. Future studies should involve a more advanced and modern equipment that is capable of automatically flying at a constant height above the seafloor.

Disparity from the constant speed assumption can only introduce errors in the computed coordinates of the extracted frames. Such a problem can easily be elevated by frequently recording the timestamp of the ROV footage with the corresponding GPS position. In this work, only the time and position of the starting and stopping locations were recorded.

The developed models can be trained on footage captured by any sensor and in any lighting conditions. Comparison of footage recorded by the down facing and front facing cameras is not straight forward because of differences in the mounting angles and fields of view. Direct matching is difficult because the maerl patches attain slightly different shapes. Although this research focused on the classification of just two taxonomy classes, planned future work includes the extraction of texture information to allow for more detailed studies and categorizations.

## Conflict of interest

The authors have no conflicts of interest to declare.

## References

Agnesi, S., Annunziatellis, A., Casese, M. L., DiNora, T., La Mesa, G., Mo, G., et al. (2009). Analysis on the coralligenous assemblages in the mediterranean sea: A review of the current state of knowledge in support of future investigations. In C. PergentMartini, & M. Brichet (Eds.), *UNEP-MAPRAC/SPA (2009) Proceedings of the 1st Mediterranean symposium on the conservation of the coralligeous and other calcareous bio-concretions*. Tunis: RAC/SPA Publication.

Borg, J. A., Lanfranco, E., Mifsud, J. R., Rizzo, M., & Schembri, P. J. (1999). Does fishing have an impact on maltese maerl grounds? Tech. rep., BIOMAERL team (1999) Final Report (in 2 Vol.). In *BIOMAERL project (Co-ordinator: P.G. Moore, University Marine Biological Station Millport, Scotland), EC Contract No. MAS3-CT95-0020*.

Boyd, D. (2009). *General purpose bombs*.. Available online: http://www.wwiiequipment.com/index.php?option=com_content&view=article&id=99:20lb-and-40lb-general-purpose-bombs&catid=43:bombs&Itemid=60 (Accessed 07.09.16).

Coiras, E., & Williams, D. (2009). Approaches to automatic seabed classification. In Peng-Yeng Yin (Ed.), *Pattern recognition*. InTech. http://dx.doi.org/10.5772/7530. Available from: http://www.intechopen.com/books/pattern-recognition/approaches-to-automatic-seabed-classification

Dimech, M., Borg, J. A., & Schembri, P. J. (2004). Report on a video survey of an offshore area off Zonqor Point (south-eastern coast of Malta), made in April 2004 as part of baseline ecological surveys in connection with the establishment of an 'aquaculture zone'. In *Report I-Preliminary video characterization [Survey commissioned by the Malta Environment and Planning Authority]*. pp. 14. Msida, Malta: Malta University Services Ltd.

Finkbeiner, M., Stevenson, B., & Seaman, R. (2001). *Guidance for benthic habitat mapping: An aerial photographic approach*. pp. 75. Charleston, SC: NOAA Coastal Services Center, National Oceanic and Atmospheric Administration. Tech. rep., Technology Planning and Management Corporation (NOAA/CSC/20117-PUB)

Hinz, H., Murray, L. G., Gell, F., Hanley, L., Horton, N., Whiteley, H., et al. (2010). *Seabed habitats around the Isle of Man. Fisheries & conservation. Tech. Rep. 12*. pp. 22. Bangor University.

Kohavi, R., & Quinlan, J. R. (2002). *Decision-tree discovery. Handbook of data mining and knowledge discovery, chapter 16.1.3*. pp. 267–276. Oxford University Press.

Landmark, K., Solberg, A. H. S., Austeng, A., & Hansen, R. E. (2014). Bayesian seabed classification using angle-dependent backscatter data from multibeam echo sounders. *IEEE Journal of Oceanic Engineering*, *39*(4), 724–739.

Mitchell, T. M. (1997). *Machine Learning* (1st Ed., pp. 432). New York, NY, USA: McGraw-Hill, Inc.

Moškon, S., Žibert, J., & Kavšek, B. (2015). Mapping of marine meadows using multibeam sonar data. *Geografski vestnik (Geographical Bulletin)*, *87*(December (1)).

Sciberras, M., Rizzo, M., Mifsud, J. R., Camilleri, K., Borg, J. A., Lanfranco, E., et al. (2009). Habitat structure and biological characteristics of a maerl bed off the northeastern coast of the Maltese Islands (central Mediterranean). *Marine Biodiversity*, *39*(4), 251–264.

Stephens, D., & Diesing, M. (2014). A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data. *PLOS ONE*, *9*(4), e93950. http://dx.doi.org/10.1371/journal.pone.0093950